

SYSTEM AND METHOD FOR MANAGING AND PRESENTING INFORMATION DERIVED FROM GENE EXPRESSION PROFILING

Publication number: JP2003505749T

Publication date: 2003-02-12

Inventor:

Applicant:

Classification:

- international: **G01N33/48; G06F17/30; G06F19/00; G01N33/48;
G06F17/30; G06F19/00; (IPC1-7): G06F17/30;
G01N33/48**

- european: **G06F19/00A3**

Application number: JP20000601434T 20000222

Priority number(s): US19990121432P 19990223; WO2000US04338
20000222

Also published as:



WO0050889 (A1)



EP1171766 (A1)



MXPA01008505 (A)



EP1171766 (A0)



CA2362544 (A1)

more >>

Report a data error here

Abstract not available for JP2003505749T

Abstract of corresponding document: **WO0050889**

A data processing system and method for managing and presenting information derived from a differential expression of genetic information which can be used to model a physiological response of biological cells. A data-represented map of metabolic pathways is provided. The map has a matrix of regions and locations within each of the regions. The regions are defined by respective coordinate sets. Each of the metabolic pathways include, at certain ones of the locations, respective graphical representations of a reaction, a substrate and a product of the reaction, and at least one biological catalyst of the reaction. An association mechanism links previously generated sets of affected sequence data to biological catalyst identifiers, each identifying one or more unique biological catalyts, and which links the biological catalyst identifiers to the coordinate sets. The sets of affected sequence data represent a direction and a magnitude of regulation of each one of a number of different nucleic acid sequences.

Data supplied from the *esp@cenet* database - Worldwide

(19) 日本国特許庁 (J P)

(12) 公表特許公報 (A)

(11) 特許出願公表番号

特表2003-505749

(P2003-505749A)

(43) 公表日 平成15年2月12日 (2003.2.12)

| | | | |
|---------------------------|-------|---------------|------------------------|
| (51) Int.Cl. ⁷ | 識別記号 | F I | テコード [*] (参考) |
| G 0 6 F 17/30 | 1 7 0 | G 0 6 F 17/30 | 1 7 0 F 2 G 0 4 5 |
| G 0 1 N 33/48 | | G 0 1 N 33/48 | Z 5 B 0 7 5 |

審査請求 未請求 予備審査請求 有 (全 41 頁)

| | | | |
|---------------|-----------------------------|-----------|--|
| (21) 出願番号 | 特願2000-601434(P2000-601434) | (71) 出願人 | ワーナー・ランバート カンパニー アメリカ合衆国, ミシガン 48105, アン アーバー, プリマス ロード 2800 |
| (96) (22) 出願日 | 平成12年2月22日 (2000.2.22) | (72) 発明者 | ロジャース, ジョン, シー アメリカ合衆国 ミシガン, マンチェスタ ー, プレザント レイク ロード 11615 |
| (85) 翻訳文提出日 | 平成13年8月21日 (2001.8.21) | (74) 代理人 | 弁理士 浅村 皓 (外3名) |
| (86) 国際出願番号 | P C T / U S 0 0 / 0 4 3 3 8 | Fターム (参考) | 2G045 AA35 DA13 JA03 JA04 JA20 5B075 ND20 NR12 UU26 |
| (87) 国際公開番号 | W O 0 0 / 0 5 0 8 8 9 | | |
| (87) 国際公開日 | 平成12年8月31日 (2000.8.31) | | |
| (31) 優先権主張番号 | 6 0 / 1 2 1 , 4 3 2 | | |
| (32) 優先日 | 平成11年2月23日 (1999.2.23) | | |
| (33) 優先権主張国 | 米国 (U S) | | |

(54) 【発明の名称】 遺伝子発現プロファイリングに由来する情報の管理と提示のためのシステムと方法

(57) 【要約】

生物細胞の生理学的反応をモデル化するために用いることができる遺伝情報の差別的な発現に由来する情報を管理し、提示するためのデータプロセッシング系及び方法。代謝経路のデータ表示マップを用意する。このマップは領域のマトリックスと各領域内の位置とを有する。これらの領域はそれぞれの座標セットによって定義される。代謝経路の各々は前記位置のうちの一定の位置において、反応と、反応の基質及び生成物と、該反応の少なくとも1つの生物学的触媒との各々のグラフィカル表現を包含する。結合機構が、予め生成された、影響を受ける配列データセットを、それぞれが1つ以上の特有の生物学的触媒を同定する生物学的触媒アイデンティファイヤーにリンクし、前記生物学的触媒アイデンティファイヤーを前記座標セットにリンクする。影響を受ける配列データセットが多くの異なる核酸配列の各自のレギュレーションの方向と大きさを表示する。

【特許請求の範囲】

【請求項1】 生物細胞の生理学的反応をモデル化するために用いることができる遺伝情報の差別的な発現に由来する情報を管理し、提示するためのデータプロセッシングシステムであって、

それぞれの座標セットによって定義される領域のマトリックスと各前記領域内の位置とを有する代謝経路のデータ表示マップであって、前記代謝経路の各々が前記位置のうちのある一定の位置において、反応と、反応の基質及び生成物と、該反応の少なくとも1つの生物学的触媒との各々のグラフィカル表現を含む前記データ表示マップ；及び

予め生成された、影響を受ける配列データセットを、それぞれが1つ以上の特有の生物学的触媒を同定する生物学的触媒アイデンティファイヤーにリンクし、前記生物学的触媒アイデンティファイヤーを前記座標セットにリンクする結合機構であって、前記影響を受ける配列データセットが多くの異なる核酸配列の各自のレギュレーションの方向と大きさを表示する前記結合機構を含む前記データプロセッシングシステム。

【請求項2】 前記結合機構によって対応領域にリンクされた、多くの前記影響を受ける配列データセットに従って、前記対応領域に対応等級を割り当てる等級付け機構をさらに含む、請求項1記載のデータプロセッシングシステム。

【請求項3】 前記マップの概観と前記マップの詳細図とをディスプレイするためのディスプレイ機構をさらに含む、前記概観が前記マップの前記領域と前記対応等級とをグラフィカルに表示し、前記詳細図が、前記パイオグラフィカル触媒のうちの前記影響を受ける配列データセットに対応するパイオグラフィカル触媒の位置の、影響を受ける生物学的触媒指標を含む前記代謝経路の詳細なグラフィカル表現を含み、前記影響を受ける生物学的触媒指標の各々がそれに対応する影響を受ける配列データセットのレギュレーションの方向を表示する、請求項2記載のデータプロセッシングシステム。

【請求項4】 生物細胞の生理学的反応をモデル化するために用いることができる、遺伝情報の差別的発現に由来する情報を管理し、提示する方法であって、

それぞれの座標セットによって定義される領域のマトリックスと各前記領域内の位置とを有する代謝経路のデータ表示マップであって、前記代謝経路の各々が前記位置のうちのある一定の位置において、反応と、反応の基質及び生成物と、該反応の少なくとも1つの生物学的触媒とのそれぞれのグラフィカル表現を含む前記データ表示マップを用意する工程；及び

予め生成された、影響を受ける配列データセットを、それぞれが1つ以上の特有の生物学的触媒を同定する生物学的触媒アイデンティファイヤーにリンクし、前記生物学的触媒アイデンティファイヤーを前記座標セットにリンクする工程であって、前記影響を受ける配列データセットが多くの異なる核酸配列の各自のレギュレーションの方向と大きさを表示する前記リンクング工程を含む前記方法。

【請求項5】 前記リンクングによって対応領域にリンクさせた、多くの前記影響を受ける配列データセットに従って、前記対応領域に対応等級を割り当てる工程をさらに含む、請求項4記載の方法。

【請求項6】 前記マップの概観と前記マップの詳細図とをディスプレイする工程をさらに含み、前記概観が前記マップの前記領域と前記対応等級とをグラフィカルに表示し、前記詳細図が、前記バイオグラフィカル触媒のうちの前記影響を受ける配列データセットに対応するバイオグラフィカル触媒の位置の、影響を受ける生物学的触媒指標を含む前記代謝経路の詳細なグラフィカル表現を含み、前記影響を受ける生物学的触媒指標の各々がそれに対応する影響を受ける配列データセットのレギュレーションの方向を表示する、請求項5記載の方法。

【請求項7】 データを記録させたマシンリーダブル媒体であって、データがコンピューターに読み込まれ、実行されるときに、該データがコンピューターに、

それぞれの座標セットによって定義される領域のマトリックスと各前記領域内の位置とを有する代謝経路のデータ表示マップであって、前記代謝経路の各々が前記位置のうちのある一定の位置において、反応と、反応の基質及び生成物と、該反応の少なくとも1つの生物学的触媒との各々のグラフィカル表現を含む前記データ表示マップを使用させ；

予め生成された、影響を受ける配列データセットを、前記影響を受ける配列データセットが多くの異なる核酸配列の各自のレギュレーションの方向と大きさを表示する場合に、それぞれが1つ以上の特有の生物学的触媒を同定する生物学的触媒アイデンティファイヤーにリンクさせ、前記生物学的触媒アイデンティファイヤーを前記座標セットにリンクさせるような前記マシンリーダブル媒体。

【請求項8】 前記データがさらにコンピューターに、

前記リンキングによって対応領域にリンクさせた、多くの前記影響を受ける配列データセットに従って、前記対応領域に対応等級を割り当てさせる、請求項7記載のマシンリーダブル媒体。

【請求項9】 前記データがさらにコンピューターに、

前記マップの概観と前記マップの詳細図とをディスプレイさせ、前記概観が前記マップの前記領域と前記対応等級とをグラフィカルに表示し、前記詳細図が、前記バイオグラフィカル触媒のうちの前記影響を受ける配列データセットに対応するバイオグラフィカル触媒の位置の、影響を受ける生物学的触媒指標を含む前記代謝経路の詳細なグラフィカル表現を含み、前記影響を受ける生物学的触媒指標の各々がそれに対応する影響を受ける配列データセットのレギュレーションの方向を表示する、請求項8記載のマシンリーダブル媒体。

【発明の詳細な説明】

【0001】

(発明の背景)

1. 著作権表示

この特許書類のある一定の部分は著作権保護を受けることができる。この特許書類はU. S. Patent and Trade Officeの特許ファイル及び記録に存在するので、この特許書類の如何なる者によるファクシミリ複写も許可されず、他の使用又は複写も許可されず、著作権所有者はあらゆる著作権利を如何なるものであっても保有する。

【0002】

2. 発明の分野

本発明は、核酸配列の差別的発現、例えば遺伝子発現プロファイリングをモニターするための方法に由来する情報を管理し、提示するための一定のシステム及び方法に関する。

【0003】

3. 背景情報の説明

遺伝子発現プロファイリングプロセスは、特定の化合物、処置又は疾患に対する細胞の生理学的反応を表示するために一般に用いられる。例えば、Iyer等による1999年1月1日の論文、Science, 283巻, 83-87頁(www.sciencemag.org)は、処置に対するヒト細胞の生理学的反応—特に、血清に対する線維芽細胞の反応を表示するための遺伝子発現の時間的プログラムの使用を開示する。cDNAマイクロアレイを用いて、8,600を超える異なるヒト遺伝子を表示している。ヒト新生児包皮から培養された、線維芽細胞を、血清から細胞を48時間取り出すことによって、静止状態にした。次に、10%FBS含有培地を加えることによって、これらの線維芽細胞を刺激し、次にマイクロアレイを用いて、異なる12時点において8,613種類のmRNA配列のレベルを測定した。このマイクロアレイを用いて、実質的に抑制されたか、又は誘発されたような遺伝子(発現配列タグ—ESTsを包含)を同定し、抑制度又は誘発度(即ち、フォールド(fold)変化)を同定した。処置に

反応してmRNAレベルが変化した517遺伝子を選択して、ヒエラルキー(hierarchy)に従ってグラフィカルに表現した。

【0004】

この情報から、種々なタンパク質を同定し、それらの生物学的機能に従って分類することができる。このような同定された生物学的機能カテゴリーはシグナル伝達、中間体-早期転写因子、他の転写因子、細胞サイクルと増殖、コアギュレーションとホメオスタシス、炎症、血管新生、組織リモデリング(tissue remodeling)、細胞骨格改造、再上皮化、コレステロール生合成及び、創傷治癒における同定されていない役割りであった。

【0005】

種々なテクノロジーが多数の遺伝子を発現させるために利用可能である。このようなテクノロジーを組み入れた有効な実施の小さい例は、SAGE(遺伝子発現の連続的分析)、オリゴアレイ及びcDNAアレイを包含する。

【0006】

このようなテクノロジーは多数の遺伝子と、それらの抑制又は誘発度とを同定するデータを生じる。これらの大きなデータセットの分析を助成するために、生物学的コンピューター分析システムが開発中である。それぞれのアレイを含む対照(control)及び処置プロープを作製するために典型的に用いられるアプローチは、Iyer等によって用いられたものであり、このアプローチでは、アップ又はダウンレギュレートされる(誘発又は抑制される)遺伝子クラスターの分布を示す、このようなアレイからのデータが二次元クラスターイメージの形態で提示される。

【0007】

処置データの研究を容易にするデータベースとウォールチャートとが提供されている。例えば、Zymed Laboratoriesによって分配される、Boehringer-Mannheim生化学経路ウォールチャートとCell Signaling Pathways Chartとは、実際に存在する選択的な代謝経路をグラフィカルに例示し、例示された種々な代謝経路間の相互関係(例えば、代謝経路間の結合と、基質代謝の分岐点)と、一定の代謝経路内

での1点から他の点までのターンオーバーの方向と速度を制御するファクターとを例示する。

【0008】

重要な代謝経路情報と、作用機構と、薬物不活化及びクリアランスの機構と、可能な副作用とをより容易に同定するために、発現プロファイリング方法から得られたデータの分析を良好に促進するシステムが必要とされている。このようなシステムは好ましくは、ある一定の処置の生理学的影響と、影響を受ける代謝活性に関連した生物学的機能との同定を助成する有意義な情報をも提供する。

【0009】

4. 定義

本明細書に開示された本発明と実施態様との、解明のためと、読者の理解を助けるために、本明細書に用いる多くの用語を次のように定義する：

【0010】

生物学的機能：

ある一定の遺伝子、タンパク質、核酸配列又は経路についての推定された機能的分類。生物学的機能の幾つかの例は代謝、血管新生、シグナル伝達、転写因子、細胞サイクル制御、増殖のレギュレーション、コアギュレーションとホメオスタシス、炎症及びアポトーシスである。

【0011】

酵素：

生化学的反応を触媒するタンパク質。

タンパク質分子：

アミノ酸の1つ又は幾つかのポリペプチド鎖。

【0012】

発現プロファイリング：

細胞由来サンプル中のある一定の核酸配列（例えば、mRNAs、タンパク質、遺伝子、ESTs）のレベルを測定して、異なるサンプルからの又は異なる時点における同じサンプルからの同じ核酸配列のレベルに対して比較するために、遺伝子発現方法を用いるプロセス。

【0013】

遺伝子：

特定のポリペプチド鎖を特定するヌクレオチド配列。

代謝経路：

基質と、反応によって生じる生成物並びにこのような反応の触媒とを包含する任意の個々の生物学的反応。代謝経路における反応の触媒は典型的に酵素である。代謝経路はまた、このような個々の反応の任意の関連シリーズをも包含する。

【0014】

作用機構：

例えば、どの特異的な、又は個体内のどこで、代謝経路若しくは生物学的機能が複合行為又は処置行為(a compound or treatment act)を行って、ある一定の生理学的効果を生じるかを同定する、バリエーション(variant)と、該バリエーションに対する反応との間の因果関係(causal link)。例えば、血圧を下げる場合に、作用機構は、特異的な代謝経路と生物学的機能とが血圧低下に作用する又は関与することを含む。

【0015】

mRNA (メッセンジャーRNA)：

DNA鋳型から一酵素RNAポリメラーゼによって合成されるRNA分子。mRNAはポリペプチド鎖のアセンブリ、翻訳として知られたプロセスのための鋳型として機能する。

生理学的影響：

何らかの生理学的変化又は反応。生理学的影響はある一定の生物学的系の状態(活性化又は不活化)、例えば高血圧における変化でありうる。

【0016】

RNA：

リボ核酸。

RNAポリメラーゼ：

鋳型としてDNAを用いてRNAを合成する酵素。

転写：

RNA分子が酵素RNAポリメラーゼによって、鋳型としてDNAを用いて合成されるプロセス。

【0017】

(発明の概要)

したがって、上記を考慮すると、本発明は、種々な態様(aspects)及び／又は実施態様の1つ以上によって、例えば上述したような、1つ以上の目的及び利点を達成するために提示される。

本発明の目的は、発現プロファイリングに基づく有意義な情報のディスプレイを容易にするための改良された機構を提供する、このような情報は処置、化合物又は疾患に関与する生物学的機能の決定、代謝経路の同定及び作用機構の同定を容易にする。本発明の他の目的は、ディスプレイされたマップのある一定の代謝経路特徴に従って発現プロファイルデータをグループ化するように、データマイニング(data mining)を可能にするために情報を組織化し、ディスプレイする構造体を提供することである。

【0018】

それ故、本発明は、生物細胞の生理学的反応をモデル化するために用いることができる遺伝情報の差別的発現に由来する情報を管理し、提示するためのシステム又は方法、又はそれらの1つ以上の構成要素に関する。このシステムは発現プロファイリング・サブシステムを含む。発現プロファイリング・サブシステムは、細胞由来サンプルの対照及び処置セットから、多数の異なる核酸配列の各自のレギュレーションの方向と大きさを表示する、それぞれの配列データセットを生成する。核酸配列セットは、研究される生物の代謝経路マップ上の特定の領域に関連する。マップ座標の概観を与えることができ、マップの、高濃度の影響を受ける核酸配列を含むような面積又は領域はマップの他の領域から、例えば異なる色を有することによって識別されることができる。マップの、高濃度の影響を受ける核酸配列を含む領域をさらに詳細に調べて、関与する特異的代謝経路と、このような代謝経路内で影響を受ける核酸配列が果たす役割とを知ることができる。

【0019】

或いは、又はさらに、一定の代謝経路セット内の特異的な、影響を受ける核酸配列を同定するマップの概観を与えることができ、このような指標(indications)は経路セット内の阻害点を表示する第1シンボルと、散経路セット内の生物学的触媒位置を表示する第2シンボルと、例示した代謝経路セットの最終生成物の位置を表示する第3シンボルとを包含する。

【0020】

以下の詳細な説明において、本発明の非限定的な具体的実施態様としての、特に多くの図面を参照して、本発明をさらに説明する、図中で同様な参照番号は図面の幾つかの図を通して同様な部分(parts)を表示する。

【0021】

(具体的実施態様の詳細な説明)

次に図面をさらに詳細に参照すると、図1は本発明の例示された実施態様による分析システム10を示す。クライアント・コンピューター14に連結される発現プロファイリング・サブシステム12が用意される。クライアント・コンピューター14は、他の要素のなかでも特に、ブラウザ・アプリケーション16、ヒトインターフェース18及びディスプレイ20を含む。ヒトインターフェース18は、例えばキーボードとマウスとを包含する、クライアント・コンピューター14とのヒト相互作用及びクライアント・コンピューター14の制御を容易にするために、任意の標準又は他のインターフェースを含むことができる。クライアント・コンピューター14はホスト・コンピューター24に、図1にイントラネット(intranet)として例示されるネットワーク・コネクション(network connection)によって連結される。ホスト・コンピューター24はデータベース26に接続される。

【0022】

発現プロファイリングシステム12は例えばAffymetrix cDNAアレイを含むことができる。これは、細胞由来サンプルの対照及び処置セットから、多数の異なる核酸配列の各自の方向と大きさを表示する、それぞれの配列データセットを生成する。

【0023】

クライアント・コンピュータ14は、ヒトインターフェース18、ディスプレイ20及びブラウザ・アプリケーション16と共に、使用者が分析システム10を操作することを可能にする。クライアント・コンピュータ14は、インターネット22とホスト・コンピュータ24とを介してデータベース26に連絡する。

発現プロファイリング・サブシステム12は発現ファイリングデータを得て、このデータを組織化された形式でデータベース23に記憶させる。

【0024】

ホスト・コンピュータ24は、他の要素のなかでも、特に、発現プロファイリングに関連したある一定の分析プロセス工程を実施し、この発現プロファイリングから得られたデータを管理するための分析アプリケーション27を備えて用意される。データベース照会と応答を取り扱い、このデータベース照会と応答に作用するために、データベース・サーバー・ソフトウェア構成要素28が用意される。

【0025】

図2は一般に、例示した実施態様による発現プロファイリング・プロセスを示す。最初の工程52では、細胞の基底サンプル（さもなくば、対照サンプルと呼ばれる）に基づいて配列が生成される。処置済みサンプル、即ち、疾患状態になった又は特定の化合物によって処置されたような細胞に基づく細胞サンプルに基づいて、1つ以上の区別される配列を生成することができる。それぞれの工程52と56の各々を実施した後に、発現された配列タグ（ESTs）を包含する同様な配列と一緒にグループ化されるクラスター・アルゴリズム54と58が実現される。遺伝子発現アレイは典型的に完全な遺伝子配列又は完全なmRNAのみを同定するのではなく、完全な配列の短いピースを含むESTsをも同定するので、遺伝子配列ピースのクラスター化は遺剰物の排除を可能にする。一定のクラスター内の配列ピースの総数は、特定の配列を有する抑制又は誘発された遺伝子の総数を表示すると見なすことができる。

【0026】

代替クラスター化方法は、表示データを用いて、発現パターンによって、即ち

、時間の経過にわたって同様なプロファイルによってクラスター化することである。このアプローチは、例えば Iyer 等によって上記論文において行われたように、未知機能の同定を助成するために、既知機能を有する遺伝子と未知機能を有する遺伝子との比較を可能にすると考えられる。

【0027】

遺伝子クラスターが実質的に影響を受けている（即ち、抑制されているか又は誘発されている）かどうかを判定するために、基底サンプル中の生成された遺伝子数を、処置済みサンプル（単数又は複数）中の生成されて、各クラスターにクラスター化された遺伝子数と比較して、各処置済みサンプルに関して、遺伝子クラスターがレギュレートされたかどうかと、そのレギュレーションの程度及び方向についての指標(indication)を得る。

【0028】

さらに詳しくは、実施例として、例えばヒトゲノムに関する、数ゲノムの6,000を超える配列を検出することができる、例えば Affymetrix GeneChip（商標）プローブアレイのような、発現プロファイリングアレイを用いて、細胞サンプルを配列決定することができる。AffymetrixはGeneChip（商標）フルイディックス・ステーション(Fluidics station)を与え、これは核酸ターゲットの、プローブアレイ・カートリッジへのハイブリダイゼーションを自動化し、したがって、試薬のデリバリーと、ハイブリダイゼーションのタイミング及び温度とを制御する。各フルイディックス・ステーションは独立的に一定の時点において4プローブアレイを処理することができる。

【0029】

したがって、細胞ディッシュ・セットから、時間の経過にわたってRNAを単離することによって、各ターゲットを用意することができる。このような細胞の処置は、それらに例えば血清を加えることによって、エミュレートすることができる。予め定められた間隔で、少量の流体を除去して、細胞を静止期にして、反応時間を停止する。したがって、予め定められた量の液体（例えば、各5ml）を有する大きなターゲットセットを製造する。次にGeneChip（商標）フルイディックス・ステーションが自動的に各ターゲットをハイブリダイズする、

即ち、これは総てのRNAを抽出して、各分子に化学的タグを加えることによって該RNAを標識して、生じた液体の、プローブアレイへのデリバリーを制御して、mRNAsに関する配列決定情報の獲得を促進する。このことは予め定められた位置においてターゲットを光に暴露させ、アレイ内の種々な位置において収集される光子を測定するプローブアレイによって行われる。次に、mRNA（又はEST）量を、配列又は配列セグメントに対応する適当な位置においてプローブによって与えられる読取りのシグナル強度に基づいて確認する。

【0030】

図3は、例示した実施態様によって行われる分析プロセスのフローチャートである。第1工程520では、遺伝子発現プロファイリングが行われ、この時点において細胞由来サンプル／ターゲットの対照及び処理セットからそれぞれの配列データセットが生成され、得られたデータは多数の異なる核酸配列クラスターの各自のレギュレーションの方向と大きさに関する情報を包含する。工程520においてひと度遺伝子発現プロファイリングが行われたならば、同定された配列と、関連するレギュレーション情報とを含むデータセットD2が生成される。次に、工程522において、各配列クラスターが生物学的触媒アイデンティファイヤー（BCI）にマッチングされる（matched）。例示した実施態様では、BCIは例えばEC番号を含むことができる。EC番号は既知酵素分類系の一部である。各EC番号は、6主要サブディビジョンの1つを意味する第1数と、サブクラスを意味する第2数と、サブサブクラスを意味する第3数と、連続番号を意味する第4数とを含む。主要なECクラスは（1）オキシドレダクターゼー酸化還元反応、（2）トランスフェラーゼー基（CH₃）の転移、（3）加水分解酵素ー切断、H₂O、（4）リアーゼー脱離による切断、（5）イソメラーゼー幾何学的変化、及び（6）リガーゼーATP加水分解に連結を包含する。いくつかのサブクラスの例として、オキシドレダクターゼは次の通りである：

【0031】

1. オキシドレダクターゼ
 1. 1. CHOH供与体
 1. 1. 1. NAD⁺又はNADP⁺受容体

- 1. 1. 2. シトクロム受容体
- 1. 1. 3. 酸素受容体
- 1. 1. 5. キノン受容体
- 1. 1. 99 他の受容体

【0032】

工程524では、影響を受けた配列、即ち、有意にレギュレートされた（少なくとも2倍に）配列の各クラスターをそのクラスター、アップレギュレートされているか又はダウンレギュレートされているか（即ち、それぞれ、誘発されているか又は抑制されているか）及びレギュレーションの程度に従って分類し、さらに、コンツア・プロット(contour plot)30の一定のセル(cell)に入るレギュレートされた配列又は配列セグメント(ESTs)数を合計して、該セルに関連してセルに入れる(binned)。これは工程526において行われる。

【0033】

工程528では、詳細なマップ図の合計されたセクションをディスプレイする、これは実質的に影響を受けた配列に対応する代謝経路を包含する。図4は生化学経路マップのコンツア・プロット図30を例示して、図3に例示されたプロセスの工程524と526において行われる行為の解明を容易にする。

【0034】

さらに詳しくは、図4は生化学経路マップのコンツア・プロット図を表示する。例示した実施態様では、このマップはBoehringer Mannheimによる生化学経路ウォールチャートに対応する。このマップは、Boehringer Mannheimウォールチャートと同様であるか若しくは匹敵する生化学経路のグラフィカル表現、又は一定の経路若しくは経路内の点がマップ内の特定の座標セットと関連する、生化学経路の任意の他の適当なグラフィカル表現を含むことができる。例示した実施態様では、X方向に沿って14カラム(X1-X14)とY方向に沿って8列(Y1-Y8)を含むセルマトリックスが用意される。図4に示したコンツア・プロット図は、一定のセル内のEC番号を有する配列数が5規定範囲の1つの範囲内であるかどうかを示す。このような範囲は異なるパターンによって示され、0、1、2-3、4-5、6-7及び8-

10を包含する。例として、座標X8、X7におけるセルは、セル内に入るEC番号を有する1配列を含む。X7、Y7におけるセルはこのセル内に入るEC番号を有する7配列を含む。したがって、セルX8、Y7は第2範囲、1に入るとして例示され、セルX7、Y7は範囲、6-7に入る配列数を有するとして示される。

【0035】

図4には、一定のセルに入るEC番号を有する配列の異なる範囲を区別するためにパターンを示すが、着色スキームを用いて範囲を示すことが好ましい。例として、範囲1を紫色によって表示し、範囲2-3を緑色によって表示し、範囲4-5を黄色によって表示し、範囲6-7を橙色によって表示し、範囲8-10を赤色によって表示する。

【0036】

したがって、図4に示したコンツァ・プロットによって与えられる図は経路マップの種々な領域にわたる活性の迅速な概観図を与えることができ、黄色、橙色及び赤色を有するような領域は最大の活性を有する領域を意味する。したがって、マップのより詳細な図を調べるために活性量によって領域を選択することができる。

【0037】

図5は、X方向に X_{n-1} 、 X_n 及び X_{n+1} とY方向に Y_{n-1} 、 Y_n 及び Y_{n+1} の規定座標におけるある一定の生化学経路の種々な態様を例示する、生化学経路マップの小部分を示す。このマップは代謝経路のグラフィカル表現を含む。このようなグラフィカル表現は、例えば基質、生成物、生物学的触媒(BCs)、阻害剤、生物学的機能及び経路方向(1方向対反対方向での経路の方向と、反応がいずれの方向にも進行しうることを意味する両方向性経路(amphibolic pathway)方向とを示す特有のグラフィカル・アイデンティファイヤーを包含する)のようなアイテムの個々のグラフィック表現を含む。

【0038】

より詳しくは、図5に示すように、複数の経路方向シンボル40a-40dが図5に示すマップのセクションに用意される。例示したライン40bと40cの

各末端における矢印の使用は、経路方向が両方向性であることを意味する。

基質／生成物_{1, 2及び3}

を表示する、複数の基質／生成物シンボル42a-42cが用意される。このようなシンボルは例えば、化学反応の方向に依存して、基質又は生成物として用いられうる一定の化合物を同定する文字(text)を含みうる。例示した実施態様における生物学的触媒（単数又は複数）₁と生物学的触媒（単数又は複数）₂を含めた、特定の経路に関連した各生物学的触媒又は生物学的触媒セットを、経路方向シンボルに隣接した、それぞれの生物学的触媒シンボル44a、bと共に例示する。

【0039】

生物学的触媒シンボル44aと44bを示すために、ブロックを用意する。これらのシンボルは単に、代謝経路の場合には典型的に酵素を含む、一定の生物学的触媒の一般的名称の文字表現を含みうる。BCI（生物学的触媒インデックス）シンボル46a、46bをそのそれぞれの生物学的触媒シンボル44a、44bに隣接して備え、例示した実施態様では、単にBCIの数値表現を含む。例示した実施態様では、標準命名法を用いて阻害剤を表示する文字を単に含みうる阻害剤シンボル48a、48bによって、任意の阻害剤が表示される。

【0040】

マップのある一定の領域における代謝経路が関連する生物学的機能は生物学的機能シンボル50によって表示することができ、これは例示した実施態様では一般的な命名法を用いた生物学的機能の文字表現を含む。生物学的機能の幾つかの例は、脂肪酸酸化、カロテノイド及びケトン体を包含する。他の機能は例えば硫黄代謝及びブチリン合成を包含する。

【0041】

例示した実施態様では、1つ以上のグラフィック表現が特有の色を有して、それが表示する情報の種類を同定することができる。例えば、BCIシンボル46a、46bとして役立つ文字は緑色であることができ、生物学的触媒シンボル4

4 a、4 4 bとして役立つ文字はマゼンタ又はアクアであることができ、阻害剤シンボル8 a、4 8 bとして役立つ文字は褐色であることができ、生物学的機能シンボル5 0として役立つ文字は青色であることができる。付加的な又は代替着色スキームを用いることができる。また、色の他に又は代わりに特有のグラフィカル・パターンを用いて、検査者が、他の種類の情報に対して1種類の情報を表示するものとしての特定のシンボルを容易に同定又は分類するのを促進することができる。図5に示した酵素は2種類の色、誘発された（アップレギュレートされた）場合の1色と、抑制された（ダウンレギュレートされた）場合の他の色とを有することができる。したがって、例示した実施態様では、生物学的触媒（単数又は複数）4 aと4 4 bがそれぞれマゼンタとアクアであり、生物学的触媒（単数又は複数）4 4 aが誘発されたが、生物学的触媒（単数又は複数）4 4 bが抑制された（ダウンレギュレートされた）ことを示す。

【0042】

例えば、図5に示すような、代謝経路マップ内の特異的シンボルに対して発現プロファイリング方法から得られた配列をマッピングすることによって、発現プロファイリングデータによって与えられた情報は、処置、疾患又は試験される細胞に与えられる化合物に付随する重要な関心に関連した情報の有意義なピースに迅速に関連付けることができる。発現プロファイリング実験の結果の可視化は、例えば生物学的機能（生物学的機能シンボル5 0によって表示）、代謝経路（マップ内の特定の座標に一定の代謝経路を形成するグラフィカル表現セットによって表示）及び作用機構（この同定は以下の実施例を用いてより完全に説明する）のような、情報の重要なピースを同定することによって可能になる。

【0043】

これは処置及び化合物の評価において顕著な利点を有することができ、例えば、作用機構、薬物不活化及びクリアランス機構、及び可能な副作用の同定を可能にする。

【0044】

図6は、関連する経路の選択された群の例示である。図6に示した関連経路は、例えば、さらに以下で説明する、図9に与えた“ビッグ ピクチャー (big pict

ure)” 図に示すマップ上の多くの同定された生物学的触媒に対応しうる。

【0045】

図6は、複数の経路（経路₁—経路₉）を含む複合経路を示す。各例示経路（経路₁—経路₉）は、1つ以上の代謝経路を、このような経路が実際に存在するように、含むことができる。これに関して、例えばBoehringer Mannheim生化学経路ウォールチャートを参照することができる。図6に示した特定の経路を調べて、作用機構、毒物学的作用及び副作用を同定することができる。

【0046】

多くの生化学経路は、異なる触媒によって触媒される、長く連続する異なる化学反応を包含する。生化学経路における第1関係工程はしばしば、フィードバック阻害と呼ばれるプロセスを通して経路の最終生成物によってレギュレートされる。代謝経路に沿った特異的酵素の阻害は阻害点前の中間化学物質のレベルを上昇させ、阻害点後の代謝産物のレベルを減少させる。

【0047】

図6に示した複合経路では、阻害点Aを示す。経路における酵素は阻害点A後に抑制されるが、他の方向における酵素は阻害点A後に誘発される。これが生じると、経路は阻害され、ある一定の最終生成物の形成を妨げ、フィードバック阻害を排除する。疾患状態又は系への薬物使用に反応した特異的酵素の誘発又は抑制を利用して、疾患又は薬物によって影響される経路を同定することができる。

【0048】

例えば、図6に示すように、薬物を動物に投与する場合には、薬物が血清コレステロールレベルを低下させることが見出され、該薬物はグラフィカルに表示される経路によって明らかにされる未知機構によって作用しうる。コレステロール生合成は主として肝臓で生ずるので、肝臓を取り出して、それからmRNAを単離することができる。発現プロファイリング方法を用いて、この阻害が多数の経路において作用する非常に多くの酵素のmRNAレベルにどのように影響するかを決定することができる。その酵素レベルが薬物処置によって顕著に影響される経路が該経路であり、同様に薬物作用機構を示唆する。

【0049】

これは、ヒドロキシメチルグルタリル-補酵素A (HMG-CoA) レダクターゼの阻害剤の場合であり、コレステロール生合成における第1工程である。この工程は図6の頂部に示す。

【0050】

経路に沿って、HMG-CoAは2反応工程(図6には特に詳しくは示さない)においてアセチル-CoAによって長鎖脂肪酸に転化される。他の方向では、HMG-CoAが経路によって5炭素イソプレノイドに転化され、次に経路によって10炭素ゲラニルに転化される。他の経路後に、生成物15炭素ファルネシルが製造される。他の経路は30炭素スクワレンを製造し、次に、これが経路によってステロイド ラノステロールに転化される。次に、複数の他の反応工程を含む経路後に、コレステロールが製造される。

【0051】

薬物(HMG-CoAレダクターゼ阻害剤)を肝臓に与え、この処置済み肝臓に対して発現プロファイリングを行う場合には、HMG-CoAレダクターゼと、脂肪酸代謝(経路1-経路3の方向に沿って進行する)に関与する酵素とが誘発され、コレステロール形成に関与する酵素が抑制される。

【0052】

薬物代謝と排泄の経路の同定は同様に行われる。大抵の薬物は、排泄のために腎臓において認識される糖又は他の分子へのコンジュゲーションよりも、より大きく反応性の種へと酸化によって代謝される。この酸化工程は、シトクロムP450酵素を含めた、200を超える酵素のうちの1つ以上によって触媒され、その後に、肝臓中のコンジュゲーティング酵素によるコンジュゲーションが行われる。これらの酵素は薬物によって直接誘発されうるか、又は薬物は正常な基質と競合するので、この場合には正常な生成物は酵素経路によってあまり製造されず、この生成物によるフィードバックは減少する。

【0053】

幾つかの遺伝子の誘発は有害な効果を暗示する。薬物代謝に関与する多様な酵素は腫瘍細胞(P451 4FI)中で誘発され、薬物による誘発は、薬物が可

能性として腫瘍形成性であることを暗示しうる。さらに、薬物の代謝は有害な代謝産物を生成する可能性があり、過酸化及びタンパク分解カスケードを誘発する可能性があり、このことは薬物又は薬物代謝産物が細胞死又は損傷を生じていることを示すことができる。

【0054】

図7は一般的に、ブロック図に、図1に例示したデータベース26の構造を示す。データベース26は、他の要素のなかでも、図1に例示したような7つの表、表1（実験）、表2（データ）、表3（配列）、表4（BCIリンク）、表5（BCI数）、表6（マップリンク）及び表7（座標）を含む。

【0055】

実験、表1には、何らかの時点における発現プロファイリング・サブシステム12が入る。これは実験アイデンティファイヤー（ExpID）と、関連実験名称及び実験条件を包含する。表2は実験から得られたデータを包含し、実験同定（ExpID）、配列同定、配列ID及び、影響を受けたと同定されている各配列のフォールド変化を包含する。表1は、表2と、変数ExpIDによってリンクされる。表2は関連配列IDと、各ExpID値に関連したフォールド変化値とを有する。表2内の配列ID値は、配列表として役立つ表3の対応インデックス付き配列IDと関連する。各配列IDに関して、受け入れ変数と配列の説明を含めた、付加的な変数がそれに付随する。

【0056】

その配列IDインデックスによって表2及び表3にリンクされるBCIリンク表4が用意される。BCIリンク表4は、BCIID、配列／リンク値及びリンクスコアを含めた、各配列ID値を付随する。各BCIIDは、表5にリストされる関連BCI数（BCI）を有する。表4及び表5の各BCIIDはマップリンク表6に与えられるBCIIDインデックスにリンクされる。各BCIIDは、それに関連した座標IDを有し、これはマップリンク表6内に与えられる。マップリンク表6は座標表7に、座標ID値によってリンクされる。座標表7は生化学経路マップのX座標、生化学経路マップのY座標、及び対応するXとY座標毎のマップ上の一定の位置に関連した生物学的機能を含めた、各座標

ID値に関連した値を与える。例示した実施態様において、関係するデータベースの第3標準形に従って、データベース26を作り上げることができる。実際のデータの大部分は表1、表3、表5及び表7に記憶され、データベース中の冗長度を最初に最小にするために、リンク表、表2、表4及び表6が与えられる。

【0057】

リンク表、表2、表4及び表6は多対多関係を容易にする。このような関係は実験と遺伝子との間に存在する—多くの遺伝子が一定の実験において影響され、多くの実験が各遺伝子によって行われうる。遺伝子とBCI数（例えば、EC番号）との間にも多対多関係が存在する。例えば、多機能遺伝子は多くのEC番号を有することができ、多くの同様な遺伝子が同じEC番号を有することができる。BCI数とマップ作成座標(mapped coordinates)との間にも多対多関係が存在する。例えば、BCI数がEC番号を含み、マップがBoehringer Mannheim生化学経路ウォールチャートを含むか、又はこれに従ってモデル化される場合に、1つのEC番号が1つの座標内又は多重座標中に容易に再三出現することができ、各座標が多くのEC番号を有することができる。

【0058】

図8は、分析系10によってデータベース26のその使用に関して行われる、データ取り扱いプロセスを例示するフローチャートである。第1工程540では、実験データを読取って、表1に記憶する。次に、工程542において、表3に配列データを記憶する行為が行われる。表1に記憶された実験データは、他のデータのなかでも、実験ID(ExpID)、実験名称(ExpName)及び実験条件を包含する。表3に記憶される配列データは配列ID、配列に対応する受け入れ番号及び、配列に関する説明データを包含する。工程544では、配列毎の(又は配列クラスター毎の)フォールド変化が決定され、この情報は表2に記憶され、ExpID及び配列IDを包含する他のデータに関連付けられる。工程546では、BCIsが配列にリンクされる。次に、表4を用いて、配列を表5のBCIsデータにリンクすることができる。

【0059】

工程548では、BCIsをマップのマップ座標にリンクする。リンク表の表

6は、表7の座標データにBCIsをリンクするために用いられる。

【0060】

図9はマップの他の概観ディスプレイを示す。この図では、阻害点60を第1シンボル60（例示した実施態様では方形である）によって、経路における阻害が生じる点に相当するマップの特定のセル内の特異的位置にディスプレイする。第2シンボル62a-621は、処置によって影響を受ける配列に対応する酵素を表示する。1つの色（図9における暗灰色）を用いて、誘発される酵素を表示し、他の色（図9では白色）が抑制された酵素を表示する。第3シンボル64aと64bは例示した経路の最終生成物を表示する。図9に示したシンボルは総て共通の複合経路上に存在する。最終生成物シンボル64aは、誘発された酵素に対応する経路の最終生成物であるので、暗灰色で示され、最終生成物シンボル64bは、抑制された酵素が存在する経路に対応する最終生成物であるので、白色として示される。

【0061】

コンツァ・マップ図が図4に示すように与えられる第1ディスプレイ・モードと、それぞれの概観経路(overview pathway)が図9に示すように与えられる第2ディスプレイ・モードとを含めた、種々なディスプレイ・モードが与えられるように、分析アプリケーション27を形成することができる。第2モードであるときに、各複合経路を別々に単独で例示することができる、又は関連の無い複合経路が総て示される1つのマップを用意することもできる。

【0062】

マップの詳細図が備えられる第3ディスプレイを用意することができる。このモードは、使用者が第1及び第2ディスプレイ・モードのいずれであっても概観ディスプレイ中の所望の座標上を単にクリックして、任意の所望の座標セットにおける詳細マップを選択的に選ぶことによって、開始することができる。

【0063】

図10は、図9に示した概観ディスプレイを作成するために、分析アプリケーション27によって行われるような工程の流れ図である。第1工程550では、BCIsの特異的座標を決定する行為が行われる。次の工程552では、同じ経

路に共通であるBCIsが決定される。2つ以上(more than one)の別々の関連の無い複合経路が存在する場合には、複数のBCIsセットが決定され、別々に分類される。工程554では、特定の共通経路の誘発されたBCIsがディスプレイされ、1つの色が誘発されたBCIsを表示し、他の色が抑制されたBCIsを表示する。

【0064】

工程556では、阻害点の亜座標(subcoordinate)が決定される。阻害点が存在する場合、即ち、共通経路の一方の側が総ての抑制BCIsを包含し、共通経路の他方の側が総ての誘発BCIsを包含する場合。この点は生化学経路マップ内の適当な位置に第2シンボルによってディスプレイされる。

【0065】

工程558では、共通複合経路の最終生成物の亜座標を決定し、このような点を第3シンボルによってディスプレイする、1つの色が誘発BCIsに対応する経路部分の最終生成物を表示し、他の色が抑制BCIsに対応する経路部分の末端に相当する最終生成物を表示する。

【0066】

阻害点は、例えば、酵素が1つの影響を受けた状態（例えば、誘発）から他の状態（例えば、抑制）へ転換する、経路に沿った点を同定することによって決定することができる。例えば、酵素がもはや影響を受けない、経路に沿った点を決定することによって、又は該当経路について知られたデータを用いて、最終生成物を推定することができる。

【0067】

次に、本発明の他のより詳しい実施態様を説明する。この実施態様は例示的な実施例であるに過ぎない。

最初に、EC（酵素コミッション）番号をBoehringer Mannheim生化学経路ウォールチャート上の座標に関連付けるデータベースを作成する。このデータベースは総てのEC番号に関する現在の説明と、EC番号に関する他の情報とを含有する。EC番号と他の酵素データとの説明は公衆に入手可能であり、ウェブサイト <http://www.expasy.ch/txt/e>

nzyme. get. から得ることができる。次に、EC番号をBoehringer Mannheim生化学経路ウォールチャートに対応する特定のマップ座標にリンクして、データベースを作成することができる。

【0068】

発現プロファイリングを行って、実験データを得たならば、EC番号を実験で得られた配列クラスターに割り当てる。これは、プロファイリング実験セットにおいて2倍を超えて(more than two fold)影響されたような、影響を受けた遺伝子に対応するGenBank受け入れ番号のリストを包含することができる。GenBank記録は<http://www.ncbi.nlm.nih.gov/entrez/>で入手可能であり、EC番号の数字のパターン(＃．＃．＃．＃)に関して解明されることができる。GenBankファイル中のEC番号の総ての出現に関して、データベース中への装填のために、GenBank受け入れ番号と対応EC番号とをテキスト・ファイルに書き加えることができる。下記はGenBankファイルのサンプルである：

【0069】

場所 4191746 375 aa 27-JAN-1999
 定義 アルコールデヒドロゲナーゼ ADH.ACCESSION 4191746PID g4191746
 データベース GENBANK 場所 L30113, アクセス L30113KEYWORDS
 ソース baboon. 生物 Papio hamadryas
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Cercopitheciidae; Cercopitheciinae; Papio.

参考文献 1 (残り1~375)

著者 Cheung,B., Holmes,R.S., Eastaugh,S.及び Beacham,I.R.
 題名 CatarrhineにおけるクラスIアルコールデヒドロゲナーゼ遺伝子の
 進化: 遺伝子変換、置換速度及び遺伝子制御
 (Evolution of Class I Alcohol Dehydrogenase Genes in Catarrhine
 Primates: Gene Conversion, Substitution Rates, and Gene Regulation)

刊行物 Mol. Biol. Evol. 16 (1), 23-36 (1999)

特徴 位置/修飾子 ソース 1..375

/生物="Papio hamadryas"

/db_xref="taxon:9557"

/組織型="腎臓" タンパク質 1..375

/注 ="ADH"

/生成物="アルコール・デヒドロゲナーゼ"

/EC_番号 ="1.1.1.1" CDS 1..375

/注="推定"

/コード化="L30113:53..1180"ORIGIN

```

1 mstgkvik kaavtvevkk pfsiceeva ppkahevrik mvavgiersd dhvsgtlv
61 plpailghea agivegvgeg vrvkpgdkv iplfpqcgk crvcknpsn ycfkndlsnp
121 rgumqdgmr ficgkpihh flgistsqy rvdnavak idaaoplekv cligcgfsig
181 ygpavkvakv tpgstcavfg lggvgisavm gckaagaari iavdinkdkf akakelgate
241 cinpqdykkp iqevlkermd ggvdfsfevi grlritmasi lccheacgs vivgvppdsq
301 nlsinpvlll tgrtwkgais ggfrksesvp klvsdfmakx fsldalinv lpfekinegf
361 dlrlsgksir tilmf//

```

【0070】

EC番号がGenBankファイルにおいて入手可能でない場合には、発現プロファイリングから得られた特定のクラスターに対応するヌクレオチド又はアミノ酸配列をGenBankファイルから得ることができ、BLAST配列アライメントを行うことができ、これはhttp://www.ncbi.nlm.

`nih.gov/cgi-bin/BLAST/nph-newblast?J`
`form=0`によって公衆に入手可能なアプリケーションにアクセスすることによって実行することができる。次に、GenBankファイルを、1未満の`e-30`である期待値（E値、BLAST結果における最も右側のカラム）とアラインする各配列に関して、これらの関連配列ファイルにおいてEC番号を探すことによって、フェッチすることができる。EC番号が存在する場合には、発現プロファイリング実験において影響を受ける遺伝子の受け入れ番号を記録することができ、配列アラインメントからの期待値を、関連配列ファイル（単数又は複数）中に見出されるEC番号（単数又は複数）と共に、同様に記録することができる。

【0071】

この時点で、この明細書において既述したように、データベースを作成することができる。これに関して、次に述べる特別な実施態様によると、図1に示するようなデータベース26はORACLEデータベースを含むことができ、ホストコンピュータはSilicon Graphics Origin 2000コンピュータを含むことができる。これらのアイテムは例示的であるに過ぎず、本発明を如何なる意味でも限定することを意図しない。他のコンピュータ、データベース及びデータベース構造を用いることができる。

【0072】

標準HTMLとPerlを用いるNetscape Fast Track WWWサーバーを用いて、分析アプリケーション27を実行することができる。このアプリケーションの実行に用いることができるPerlモジュールは（1）DBI/DBD—リモート・バスマップ・データベースに連絡するためのデータベース・インターフェース、（2）CGI—HTMLコードを生成するため、（3）PGPLOT—コンツァ・プロットを作成するためのコンパイルしたPGPLOT Fortranライブラリへのインターフェース、（4）GD—PGPLOTによって製造されたGIFイメージを回収するため及びバックグラウンド着色に用いる多角形及び長方形を描くためのグラフィカル・ドローイング・モジュール（graphical drawing module）、（5）MLDBM—イメージ

・マップ・シェイプ・データを作成するために持続的なマルチレベル・データ構造の作成を可能にするPerlモジュール、及び(6) ImageMagick- GDによって作成されるバックグラウンドを用いて、マスク、オーバーレイ及びバックグラウンド着色を生じることができるように、イメージ・プロセッシングを実施するためのモジュールを包含する。

【0073】

アプリケーションは、使用者がブラウザー・アプリケーション16を用いてバス・マップ・ウェブ・ページに接続し、実験を選択し、データベースに照会して、実験において2倍より大きく影響された遺伝子のウォール・チャート座標を選択することができるように構成することができる。各マップ座標にマッピングされた(mapped)遺伝子数を入れ、座標毎のヒット(hits)のコンツア・プロットを例えば図4に示すようにディスプレイすることができる。例えば、図9に示すような、他のディスプレイも同様に与えることができる。使用者はマウスを用いてカーサーをマップイメージ上の位置に動かして、マップのその領域に対応する生物学的機能を知ることができ、マップの特定のセル上をクリックして、例えば図5に示すような、経路情報のより詳細な図を得ることができる。これに関して、Boehringer Mannheim生化学経路ウォールチャート構造を用いる場合には、これを改変して、誘発された遺伝子及び抑制された遺伝子、並びにこれらの遺伝子に対応する同定済み酵素に関連したEC番号を例示する。影響を受ける遺伝子に対応する酵素を、この酵素が抑制されるか又は誘発されるかに基づいて、着色する。詳しくは、対応する遺伝子が誘発される場合には、酵素をマゼンタ文字で表示することができ、該遺伝子が抑制される場合には、シアンで、同じEC番号を有する2つ以上の遺伝子クラスターが対立する方向で影響された場合には、緑色で表示することができる。ブラウザー・アプリケーション16によって使用者に与えられるインターフェースはディスプレイ20上にディスプレイされ、特定の遺伝子と該遺伝子に関連する総ての利用可能な実験とに関する情報を得るために、使用者が受け入れ番号上をクリックすることを可能にする機構を与えることができる。特定のEC番号上をクリックして、そのEC番号に関するあらゆる情報を得ることを可能にする機構を与えることもできる。さらに、使

用者が任意の特定のパラメーターによって照会して、そのパラメーターに関する情報を得ることができるように、分析システム10にサーチ・ツールを備えることができる。例えば、使用者は受け入れ番号又は遺伝子説明によって照会して、重要な特定遺伝子に関する情報を見出すことができる。

【図面の簡単な説明】

【図1】

遺伝子発現プロファイリング・データ分析システムのブロック図。

【図2】

遺伝子発現プロファイリング・プロセスの流れ図。

【図3】

遺伝子発現プロファイリングに由来する情報を管理するためのプロセスの流れ図。

【図4】

マップ内の一定座標における影響を受ける核酸配列の濃度を示す、生化学経路マップの概観図。

【図5】

生化学経路マップの一定領域内のある一定セルの詳細な拡大図。

【図6】

フィードバック阻害によって影響される、ある一定の生化学経路セット。

【図7】

例示した実施態様によるデータベース構造の線図。

【図8】

発現プロファイリング・データをマッピングした代謝経路に突き合わせるためにクライアント・コンピューターによって行われるプロセスを表示するフローチャート。

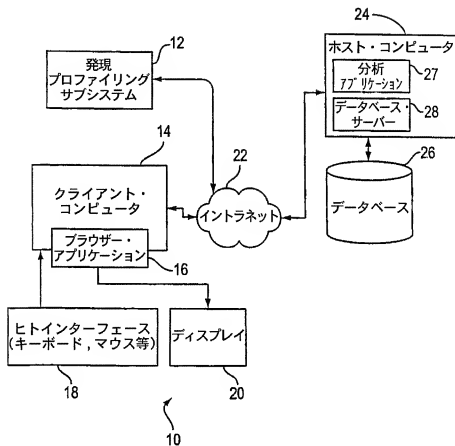
【図9】

代謝経路マップの概観ディスプレイの例、このマップには、関連する抑制された及び誘発された生物学的触媒、阻害点及び最終生成物が表示される。

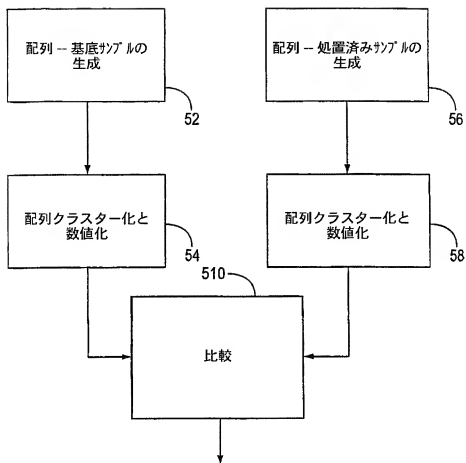
【図10】

概観マップ・ディスプレイ上の簡略シンボルセットによって、影響を受ける経路内のBCIsを同定するプロセスの流れ図。

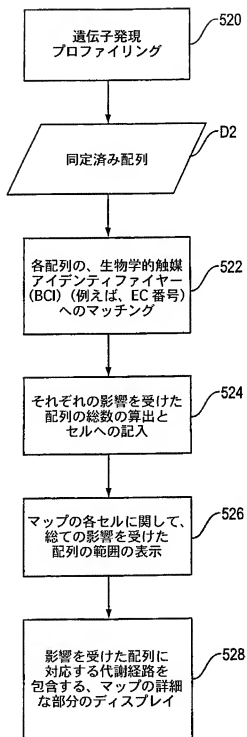
【図1】



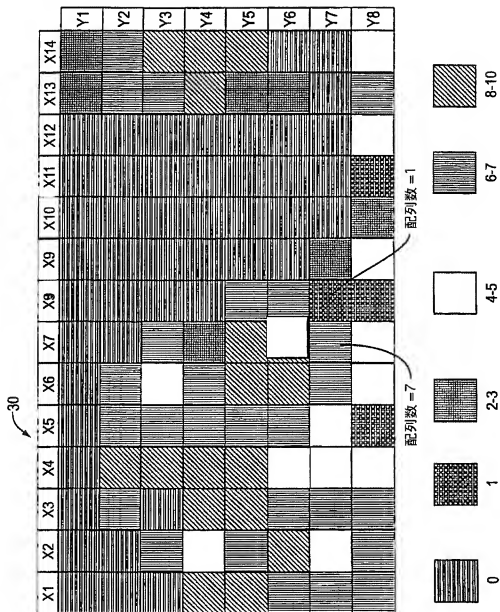
【図2】



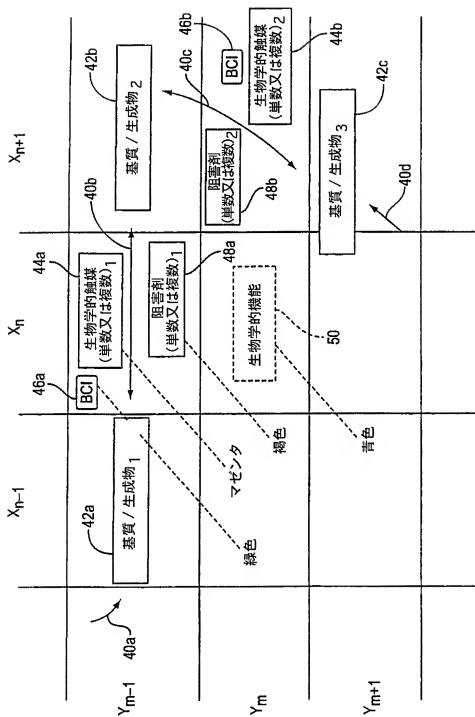
【図3】



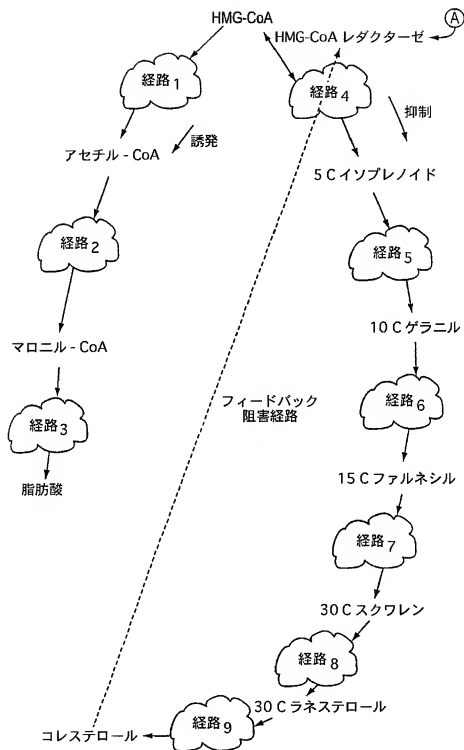
【図4】



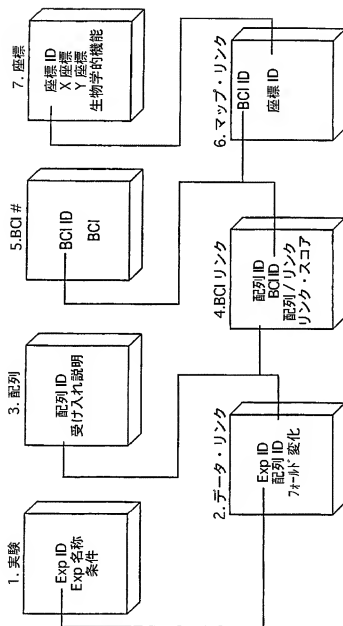
【図5】



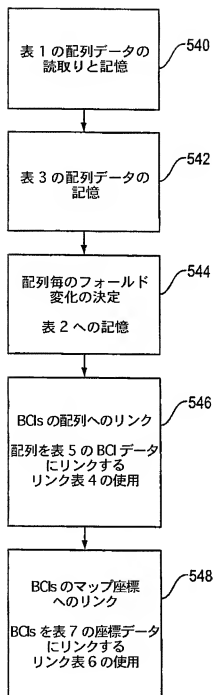
【図6】



【図7】



【図8】



【図9】

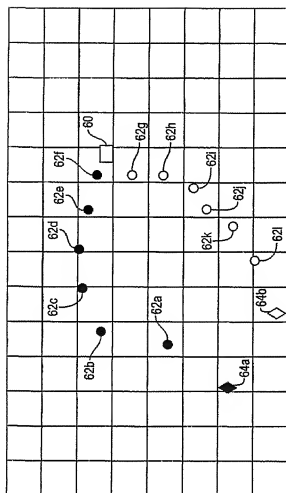
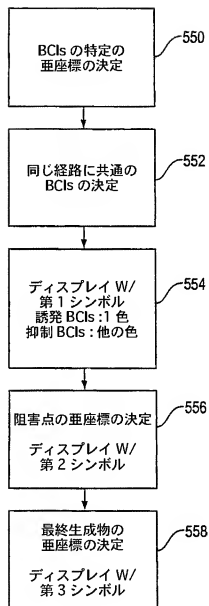


FIG. 9

【図10】



【国際調査報告】

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/US00/04338

| | |
|---|--|
| A. CLASSIFICATION OF SUBJECT MATTER IPC(C) : G01N 31/00; G06F 19/00. US CL : 702/20, 19, 27; 435/6 According to International Patent Classification (IPC) or to both national classification and IPC | |
| B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) U.S. : 702/20, 19, 27; 435/6 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EAST, WEST, Derwent, MEDLINE, CAPS, Bioscience (STN) expression profiling, map, sequence, computer/processor | |
| C. DOCUMENTS CONSIDERED TO BE RELEVANT | |
| Category* | Citation of document, with indication, where appropriate, of the relevant passages Relevant to claim No. |
| X | BASSETT JR., D.E. et al. Gene Expression Information- It's all in Your Mine. Nature Genetics. 01 January 1999, Volume 21 (Supplement) pages 51-55, see entire document. 1-9 |
| Y | DEBOUCK, C. et al. DNA Microarrays in Drug Discovery and Development. Nature Genetics. 01 January 1999, Vol. 21 (Supplement) pages 48-50, see entire document. 1-9 |
| Y | IYER, V.R. et al. The Transcriptional Program in the Response of Human Fibroblasts to Serum. Science. 01 January 1999, Vol 283, pages 83-87, see entire document. 1-9 |
| <input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex. | |
| * Special categories of cited documents: "A" Document defining the general state of the art which is not considered to be of particular relevance. "X" Document published on or after the international filing date. "Y" Document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reasons (as specified). "C" Document referring to an oral disclosure, use, exhibition or other means. "D" Document published prior to the international filing date but later than the priority date claimed. | "E" Later documents published after the international filing date or priority date and not in conflict with the application but cited to understand the principles or theory underlying the invention. "X" Document of particular relevance; the claimed invention cannot be considered novel or obvious by consideration of this document in isolation. "Y" Document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is taken alone. "C" Document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is considered with one or more other cited documents, such consideration being confined to a person skilled in the art. "D" Document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is considered with one or more other cited documents, such consideration being confined to a person skilled in the art. |
| Date of the actual completion of the international search 19 MAY 2000 | Date of mailing of the international search report 29 JUN 2000 |
| Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box 107 Washington, D.C. 20531 Facsimile No (703) 305-2230 | Authorized officer Heather Rousseau, Esq. MARY R. ZEMAN Telephone No. (703) 305-9196 |

Form PCT/ISA/216 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No
PCT/US00/04138

| C (Continue with) DOCUMENTS CONSIDERED TO BE RELEVANT | | |
|---|--|-----------------------|
| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
| Y | DUGGAN, D.J. et al. Expression Profiling Using cDNA Microarrays. Nature Genetics. 01 January 1999, Vol. 21, (Supplement) pages 10-14, see entire document. | 1-9 |
| Y | KAWAMOTO, S. et al. Expression Profiles of Active Genes in Human and Mouse Livers. Gene. 1996, Vol. 174, pages 151-158, see entire document. | 1-9 |
| A | US 5,866,330 A (KINZLER et al.) 02 February 1999, see entire document. | 1 |
| A | BOWTELL, D.D.L. Options Available- From Start to Finish- for Obtaining Expression Data by Microarray. Nature Genetics. 01 January 1999, Vol. 21 (Supplement) pages 25-32, see entire document. | 1-9 |

フロントページの続き

(81)指定国 EP(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG), AP(GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW